



## Special Issue on Content Based Multimedia Indexing

Yiannis Kompatsiaris, Guillaume Gravier

### ► To cite this version:

Yiannis Kompatsiaris, Guillaume Gravier. Special Issue on Content Based Multimedia Indexing. Multimedia Tools and Applications, 2017, 10.1007/s11042-017-5106-y . hal-01578925

**HAL Id: hal-01578925**

**<https://hal.science/hal-01578925>**

Submitted on 30 Aug 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Special Issue on Content Based Multimedia Indexing**

Content-based multimedia indexing systems aim at providing user-friendly, fast and accurate access to large multimedia repositories from the automatic analysis of multimedia content, so as to find specific content, information, gain knowledge and insight. Over the years, various tools and techniques from different fields such as database, machine learning, pattern recognition, and human computer interaction have contributed to the success of multimedia systems.

In spite of significant progress in the field and of the “deep learning revolution”, we still face situations where multimedia systems analyzing content show limits in accuracy, generality and scalability. This special issue focuses on the recent advances in content-based multimedia indexing core technology--e.g., understanding deep learning architectures, using formal knowledge and ontologies, putting the user in the loop--and in application of the latter. Interestingly, contributions address a variety of real-life applications such as human action recognition, cataract surgery videos, visual attention of people with dementia and TV program related applications.

This special issue of Multimedia Tools and Applications follows the 14th edition of the Content-Based Multimedia Indexing (CBMI) workshop, held in Bucharest in the spring of 2016. After its first edition in Toulouse, France, in 1999, CBMI has progressively evolved as the main venue to discuss progress and trends in the field. The workshop is now a well-established forum that brings together the various communities involved in all aspects of content-based multimedia indexing for information retrieval, browsing, visualization and for multimedia analytics at large. As in previous editions, the 2016 workshop received a number of submissions from which the 60% most qualified were selected, covering all aspects of the field and drawing a clear picture of recent achievements and current trends.

Extending the panorama of research established during CBMI 2016, this special issue of MTAP was designed to include some of the best contributions of the workshop along with new contributions from the community. The call for papers was widely distributed so as to get a rich overview of current trends, reaching out far beyond the workshop attendance. The program committee of CBMI 2016 was extended to ensure proper feedback from at least two reviewers for each of the 24 submissions that we received. Out of these submissions, half were extended version of CBMI contributions, including three from a special session on healthcare.

The special issue includes 14 contributions, grouped in four broad categories, the last two categories departing from classical frameworks.

The first group of three papers presents recent work in multimedia retrieval from standard and social media. Some classical problems are revisited in light of recent achievements in content description, such as the analysis of deep neural networks features for sketch-based image retrieval or graph-based approaches to combine heterogeneous modalities

for multi and cross modal retrieval. User feedback is also addressed, investigating new approaches that combine feedback at different levels of content abstraction.

The next three papers focus on classification approaches addressing various applications. Multimodal recognition methods are considered, with several features and their combination techniques being studied, in concept detection from image-text data. Long studied problems such as texture classification still attract attention and a novel micro-macro feature combination is presented, which is highly invariant and discriminative. Computational efficiency is an important requirement in specific domains and the third paper addresses this challenge with a new approach for feature extraction and encoding that obtains real-time frame rate processing in human action recognition.

CBMI 2016 included a special session on “Content-based image and multimedia analysis and indexing for healthcare” and three papers from the special session appear in this issue addressing real-time analysis of cataract surgery videos, disease detection in gastrointestinal videos and visual attention prediction for studies of attention of patients with dementia. All contributions aim at assisting clinicians either in their training or in the diagnosis procedure.

In many applications, training examples are not available and classification must be performed in an unsupervised manner. This is the case of the the task “Multimodal person discovery in broadcast TV” as part of the MediaEval 2015 benchmarking initiative, which is presented in the first paper of this category. Person’s names are discovered in an unsupervised way from media content using text overlay or speech transcripts. The second paper addresses the problem of unsupervised TV program segmentation by exploiting the common structure appearing in episodes collection.

The last three papers address a burning question in the field: How can formal knowledge, e.g., knowledge bases or ontologies, be used to improve multimedia content analysis and indexing? Classical multimedia analysis techniques rely on machine learning approaches that can hardly accommodate formal knowledge and linking multimedia content to knowledge bases remain a difficult task except in some specific conditions. Offering a smooth transition from unsupervised classification, a weakly supervised framework for multimodal entity linking leveraging vision and language is described. Fuzzy ontologies are proposed to bridge the gap between factual content description and human perception. Another contribution is on the construction and use of a taxonomy to lower the computational cost of ensembles of exemplar support vector machines.

We believe that this special issue will offer a great opportunity for readers to quickly grasp recent advances and current trends in approaches and applications in the field of content based multimedia indexing and retrieval.

The guest editors would like to thank all the authors and reviewers for their contributions, and acknowledge the editorial support of the Editors in Chief of Multimedia Tools and Applications Journal, and their Springer colleagues.

The Guest Editors,

Yiannis Kompatsiaris  
Information Technologies Institute,  
Centre for Research and Technology Hellas, Thessaloniki, Greece  
E-mail address: *ikom@iti.gr*

Guillaume Gravier  
IRISA & Inria Rennes,  
CNRS, France  
E-mail address: *guig@irisa.fr*